

**Management Report for the Facility Registry System (FRS)**

**June 30, 2000**

frs-june-chart-final.wpd

***FRS Population Progress: OVERALL SUMMARY***

**As of June 30<sup>th</sup>, 2000,**

**FRS records have been created: 34,416 ( Goal FY 2000: 50,000+)**

**Program Records Extracted: 49,372 ( TRI-1998, RMP, RCRIS-TSDs, PCS-Majors)**

**Program Records needing Manual Review/Data Quality Validation: 13,603**

***Preliminary Findings on Facility Integration***

- S** Very few of the major facilities report to the 4 national systems in a consistent or exact manner.
- S** Using the *TRI-1998*, as the base universe of facility records since it is a report of releases from various media:
  - of the 23,576 actively reporting facilities by TRI only:
  - 9% of the reporting RMP-99 facilities are using the exact facility information
  - 3% of the reporting RCRIS-TSD facilities are using the exact facility information
  - and 1% of the reporting PCS Majors report the same facility information as the TRI self reporters.
- S** Not all the national systems store the facility identification data within their own database in a disciplined and consistent manner.
  - S** Thus automated matching must be sufficiently sophisticated to identified same place facilities.
  - S** And manual data validation and review is essential part of the integration process.
- S** Application development for cross-media purposes such as Risk Assessment, Facility Enforcement Targeting, and Environmental Justice Analysis is very difficult in the current EPA database technology and data structures.

The FRS team and its partners in the Regions and States expect to see higher integration percentage once the manual process is conducted on these program records in creating new and linking to FRS records.

## FRS Population Summary by Program System from the Automated Loading (June 2000)

Date pulled to FRS	Program System	# of extracted records	# joined to existing FRS	# of FRS records created (Running Total)	# of records being manually reviewed	% of records being manually reviewed
06/05/00	TRI (1998)	23,576	30	19,731	3,939	16% of the TRI Universe
06/16/00	RMP (1999)	14,861	1,923	9,772 new (29,503)	3,166	21% of the RMP Universe
06/20/00	RCRIS - TSDs	4,035 *	735	1,578 new (31,081)	1,722	43% of the RCRIS universe
06/27/00	PCS - Majors	6,900	1,362	762 new (31,843)	4,776	69% of the PCS Majors universe
<b>Total 06/30/00</b>		<b>49,372</b>	<b>4,050</b>	<b>34,416**</b>	<b>13,603</b>	

\*Note: 485 more RCRIS TSDs records will be extracted in July due to a failure in the extract routine.

\*\* Further reconciliation by the SDC - FRS Support Team of the sub-set of TRI, RMP and RCRIS and PCS records by manual review into the FRS includes: **2,573 new FRS records**

## Data Quality by Program System Awaiting FRS Manual Loading (June 2000)

Program System	Missing or Irregular Address/ Direction/ PO Box for Address	Invalid city, county, zip code or combination	Invalid DQ combinations/ Potential match but data discrepancies/Data gaps in Critical locational fields	Valid handler # in TRI, but the name and address do not match for computer matching
TRI (1998)	3,470	469	-	-
RMP (1999)	2,104	1,062	-	-
RCRA TSDs	770	-	345	607
PCS - Majors	1,344	530	2,333	569
<b>Total</b>	<b>7,688</b>	<b>2,061</b>	<b>2,678</b>	<b>1,176</b>

### **FINAL FRS COUNT FOR JUNE:**

As of June 30<sup>th</sup>, the Total Universe of FRS is : **34,416**

## **Locational Data Quality prior to the FRS Loading and Reconciliation (June 2000)**

<b>Program System</b>	<b># of records that do not have a Lat/Long locational coordinate</b>	<b># of records that have a Zip Code Centroid</b>
TRI (1998)	77	464
TRI (1997)	7	42
RMP (1999)	864	73
RCRIS- TSD	51	233
PCS - Majors	269	28
<b>Total</b>	<b>1,268</b>	<b>840</b>

408,041 facility linked records still have no locational data.

Of these 408, 041 facilities, 149, 693 unique facilities in FLA do not still have any locational value that are mapped for the 7 national systems of RCRA, BRS, RMP, CERCLIS, PCS, AFS, and TRI.

Of the 408,041 facilities, 258,348 unique facilities in FLA do not have a locational value that are mapped from the national systems: SDWIS, SSTS, PADS, NCDB and DOCKET.

### **NEXT STEPS**

#### **In Population:**

The next program universe into FRS in July is the TRI-97 and BRS 1997 and BRS 1995 ( hazardous waste handlers) universe. State data transfer protocols will be created and transmitted to selected states for review and discussion.

#### **In Program Linkage and Integration:**

Further review of the TRI program numbers reported in 1998 collection will be evaluated to find greater number of linkages between TRI- RCRA and PCS.

#### **In Data Quality Analysis:**

Engagement with the National System Managers and the Regional Data Stewards will be initiated in the latter part of July to undertake a data quality enhancement of records needing manual review and of the created FRS records.